Artificial intelligence Artificial intelligence (AI) is the capability of computational systems to perform tasks typically associated with human intelligence, such as learning, reasoning, problem-solving, perception, and decision-making. It is a field of research in computer science that develops and studies methods and software that enable machines to perceive their environment and use learning and intelligence to take actions that maximize their chances of achieving defined goals. [1] High-profile applications of AI include advanced web search engines (e.g., Google Search); recommendation systems (used by YouTube, Amazon, and Netflix); virtual assistants (e.g., Google Assistant, Siri, and Alexa); autonomous vehicles (e.g., Waymo); generative and creative tools (e.g., ChatGPT and AI art); and superhuman play and analysis in strategy games (e.g., chess and Go). However, many AI applications are not perceived as AI: "A lot of cutting edge AI has filtered into general applications, often without being called AI because once something becomes useful enough and common enough it's not labeled AI anymore ." [2][3] Various subfields of AI research are centered around particular goals and the use of particular tools. The traditional goals of AI research include learning, reasoning, knowledge representation, planning, natural language processing , perception , and support for robotics . [a] To reach these goals, AI researchers have adapted and integrated a wide range of techniques, including search and mathematical optimization, formal logic, artificial neural networks, and methods based on statistics, operations research, and economics . [b] AI also draws upon psychology , linguistics , philosophy , neuroscience , and other fields. [4] Some companies, such as OpenAI, Google DeepMind and Meta, aim to create artificial general intelligence (AGI)—AI that can complete virtually any cognitive task at least as well as a human. [5] Artificial intelligence was founded as an academic discipline in 1956, [6] and the field went through multiple cycles of optimism throughout its history , [7][8] followed by periods of disappointment and loss of funding, known as AI winters . [9][10] Funding and interest vastly increased after 2012 when graphics processing units started being used to accelerate neural networks, and deep learning outperformed previous AI techniques. [11] This growth accelerated further after 2017 with the transformer architecture . [12] In the 2020s, the period of rapid progress marked by advanced generative AI became known as the AI boom. Generative AI and its ability to create and modify content exposed several unintended consequences and harms in the present and raised ethical concerns about AI's long-term effects and potential existential risks, prompting discussions about regulatory policies to ensure the safety and benefits of the technology. The general problem of simulating (or creating) intelligence has been broken into subproblems. These consist of particular traits or capabilities that researchers expect an intelligent system to display. The traits

described below have received the most attention and cover the scope of AI research. [a] Goals An ontology represents knowledge as a set of concepts within a domain and the relationships between those concepts. Early researchers developed algorithms that imitated step-by-step reasoning that humans use when they solve puzzles or make logical deductions . [13] By the late 1980s and 1990s, methods were developed for dealing with uncertain or incomplete information, employing concepts from probability and economics . [14] Many of these algorithms are insufficient for solving large reasoning problems because they experience a "combinatorial explosion": They become exponentially slower as the problems grow. [15] Even humans rarely use the step-by-step deduction that early AI research could model. They solve most of their problems using fast, intuitive judgments. [16] Accurate and efficient reasoning is an unsolved problem. Knowledge representation and knowledge engineering [17] allow AI programs to answer questions intelligently and make deductions about real-world facts. Formal knowledge representations are used in content-based indexing and retrieval, [18] scene interpretation, [19] clinical decision support, [20] knowledge discovery (mining "interesting" and actionable inferences from large databases), [21] and other areas. [22] A knowledge base is a body of knowledge represented in a form that can be used by a program. An ontology is the set of objects, relations, concepts, and properties used by a particular domain of knowledge. [23] Knowledge bases need to represent things such as objects, properties, categories, and relations between objects; [24] situations, events, states, and time; [25] causes and effects; [26] knowledge about knowledge (what we know about what other people know); [27] default reasoning (things that humans assume are true until they are told differently and will remain true even when other facts are changing); [28] and many other aspects and domains of knowledge. Among the most difficult problems in knowledge representation are the breadth of commonsense knowledge (the set of atomic facts that the average person knows is enormous); [29] and the sub-symbolic form of most commonsense knowledge (much of what people know is not represented as "facts" or "statements" that they could express verbally). [16] There is also the difficulty of knowledge acquisition , the problem of obtaining knowledge for AI applications. [c] An "agent" is anything that perceives and takes actions in the world. A rational agent has goals or preferences and takes actions to make them happen. [d][32] In automated planning , the agent has a specific goal. [33] In automated decision-making , the agent has preferences—there are some situations it would Reasoning and problem-solving Knowledge representation Planning and decision-making prefer to be in, and some situations it is trying to avoid. The decision-making agent assigns a number to each situation (called the "utility") that measures how much the agent prefers it. For each

possible action, it can calculate the "expected utility": the utility of all possible outcomes of the action, weighted by the probability that the outcome will occur. It can then choose the action with the maximum expected utility. [34] In classical planning , the agent knows exactly what the effect of any action will be. [35] In most real-world problems, however, the agent may not be certain about the situation they are in (it is "unknown" or "unobservable") and it may not know for certain what will happen after each possible action (it is not "deterministic"). It must choose an action by making a probabilistic guess and then reassess the situation to see if the action worked. [36] In some problems, the agent's preferences may be uncertain, especially if there are other agents or humans involved. These can be learned (e.g., with inverse reinforcement learning), or the agent can seek information to improve its preferences. [37] Information value theory can be used to weigh the value of exploratory or experimental actions. [38] The space of possible future actions and situations is typically intractably large, so the agents must take actions and evaluate situations while being uncertain of what the outcome will be. A Markov decision process has a transition model that describes the probability that a particular action will change the state in a particular way and a reward function that supplies the utility of each state and the cost of each action. A policy associates a decision with each possible state. The policy could be calculated (e.g., by iteration), be heuristic , or it can be learned. [39] Game theory describes the rational behavior of multiple interacting agents and is used in AI programs that make decisions that involve other agents. [40] Machine learning is the study of programs that can improve their performance on a given task automatically. [41] It has been a part of AI from the beginning. [e] There are several kinds of machine learning. Unsupervised learning analyzes a stream of data and finds patterns and makes predictions without any other guidance. [44] Supervised learning requires labeling the training data with the expected answers, and comes in two main varieties: classification (where the program must learn to predict what category the input belongs in) and regression (where the program must deduce a numeric function based on numeric input). [45] In reinforcement learning, the agent is rewarded for good responses and punished for bad ones. The agent learns to choose responses that are classified as "good". [46] Transfer learning is when the knowledge gained from one problem is applied to a new problem. [47] Deep learning is a type of machine learning that runs inputs through biologically inspired artificial neural networks for all of these types of learning. [48] Learning Kismet, a robot head which was made in the 1990s; it is a machine that can recognize and simulate emotions. [64] Computational learning theory can assess learners by computational complexity, by sample complexity (how much data is required), or by other notions of optimization . [49] Natural language

processing (NLP) [50] allows programs to read, write and communicate in human languages such as English. Specific problems include speech recognition, speech synthesis, machine translation , information extraction , information retrieval and question answering . [51] Early work, based on Noam Chomsky's generative grammar and semantic networks, had difficulty with word-sense disambiguation [f] unless restricted to small domains called " micro-worlds " (due to the common sense knowledge problem [29]). Margaret Masterman believed that it was meaning and not grammar that was the key to understanding languages, and that thesauri and not dictionaries should be the basis of computational language structure. Modern deep learning techniques for NLP include word embedding (representing words, typically as vectors encoding their meaning), [52] transformers (a deep learning architecture using an attention mechanism), [53] and others. [54] In 2019, generative pre-trained transformer (or "GPT") language models began to generate coherent text, [55][56] and by 2023, these models were able to get human-level scores on the bar exam , SAT test, GRE test, and many other real-world applications. [57] Machine perception is the ability to use input from sensors (such as cameras, microphones, wireless signals, active lidar, sonar, radar, and tactile sensors) to deduce aspects of the world. Computer vision is the ability to analyze visual input. [58] The field includes speech recognition , [59] image classification , [60] facial recognition , object recognition , [61] object tracking , [62] and robotic perception . [63] Affective computing is a field that comprises systems that recognize, interpret, process, or simulate human feeling, emotion, and mood . [65] For example, some virtual assistants are programmed to speak conversationally or even to banter humorously; it makes them appear more sensitive to the emotional dynamics of human interaction, or to otherwise facilitate human–computer interaction. However, this tends to give naïve users an unrealistic conception of the intelligence of existing computer agents. [66] Moderate successes related to affective computing include textual sentiment analysis and, more recently, multimodal sentiment analysis, wherein AI classifies the effects displayed by a videotaped subject. [67] Natural language processing Perception Social intelligence Illustration of gradient descent for 3 different starting points; two parameters (represented by the plan coordinates) are adjusted in order to minimize the loss function (the height) A machine with artificial general intelligence should be able to solve a wide variety of problems with breadth and versatility similar to human intelligence . [68] AI research uses a wide variety of techniques to accomplish the goals above. [b] AI can solve many problems by intelligently searching through many possible solutions. [69] There are two very different kinds of search used in AI: state space search and local search. State space search searches through a tree of possible states to try to find a goal state. [70] For example, planning algorithms

search through trees of goals and subgoals, attempting to find a path to a target goal, a process called means-ends analysis . [71] Simple exhaustive searches [72] are rarely sufficient for most real-world problems: the search space (the number of places to search) quickly grows to astronomical numbers. The result is a search that is too slow or never completes. [15] " Heuristics " or "rules of thumb" can help prioritize choices that are more likely to reach a goal. [73] Adversarial search is used for game-playing programs, such as chess or Go. It searches through a tree of possible moves and countermoves, looking for a winning position. [74] Local search uses mathematical optimization to find a solution to a problem. It begins with some form of guess and refines it incrementally. [75] Gradient descent is a type of local search that optimizes a set of numerical parameters by incrementally adjusting them to minimize a loss function. Variants of gradient descent are commonly used to train neural networks , [76] through the backpropagation algorithm. Another type of local search is evolutionary computation, which aims to iteratively improve a set of candidate solutions by "mutating" and "recombining" them, selecting only the fittest to survive each generation. [77] Distributed search processes can coordinate via swarm intelligence algorithms. Two popular swarm algorithms used in search are particle swarm optimization (inspired by bird flocking) and ant colony optimization General intelligence Techniques Search and optimization State space search Local search <meta> •TheHTML <meta> tagprovides metadata or information about the HTML document. It includes information like character set, description, keywords, and other important details. <meta> tags always go inside the <head> element •It helps in defining the page's title, encoding, author, and viewport settings, etc. These tags are not visible on the web page but play a vital role in structuring and categorizing content for browsers and search engines.